An *In Vitro* Versus *In Vivo* Toxicogenomics Investigation of Pre-natal Exposures to Tobacco Smoke

Iain Perry[1], Keith Sexton[1], Zoe Prytherch[1], Jason Blum[2], Judith Zelikoff[2]* and Kelly Ann BéruBé[1]*

[1]*School of Biosciences, Cardiff University, Museum Avenue, Cardiff, CF10 3AX, UK.*

[2]*Department Environmental Medicine, NYU School of Medicine, NYU Langone Medical Centre, 57 Old Forge Road, Tuxedo, NY 10987, USA*

*Joint Corresponding Authors

Abstract

Approximately 1 million women smoke during pregnancy despite evidence demonstrating serious juvenile and/or adult diseases being linked to early-life exposure to cigarette smoke. Susceptibility could be determined by factors in previous generations, i.e. pre-natal or 'maternal' exposures to toxins. Pre-natal exposure to airborne pollutants such as mainstream cigarette smoke has been shown to induce early-life insults (i.e. gene changes) in Offspring that serve as biomarkers for disease later in life. In this investigation, we have evaluated genome-wide changes in the lungs of mouse Dams and their juvenile Offspring exposed pre-natally to mainstream cigarette smoke. An additional lung model was tested alongside the murine model, as a means to find an alternative *in vitro,* human tissue-based replacement for the use of animals in medical research. Our toxicogenomic and bioinformatic results indicated that *in utero* exposure altered the genetic patterns of the foetus that could put them at greater risk for developing a range of chronic illnesses in later-life. The genes altered in the *in vitro*, cell culture model were reflected in the murine model of pre-natal exposure to MCS. The use of alternative *in vitro* models derived from human medical waste tissues could be viable options to achieve human end-point data and to conduct research that meets the remits for scientists to undertake the 3Rs practises.

Keywords: toxicogenomics, pre-natal exposure, mainstream cigarette smoke*, in vitro*, *in vivo*

Introduction

Lung cancer was estimated to have accounted for over 14% of cancer cases in the US during 2014 and more than 27% of cancer related mortalities. Indeed, nearly 70% of incidences of lung cancer are predicted to end in mortality (Siegel *et al.* 2014). Despite the high incidence and mortality rates, lung cancer has historically received a disproportionately low share of funding for research. The US National Institute of Health calculated the US spent less than 5% of all cancer research funding on dedicated lung cancer research (NIH, 2015). The imbalance of research funding in the US is indicative of a wider global trend, which is often attributed to the wider social stigma that lung cancer is a direct result of smoking and the consequence of their conscious transgressions. Indeed, the complex mixture of over 4000 cigarette smoke compounds, of which many either direct or second-hand, are known to have direct links to, cancer, cell irritation and death (Ng *et al.* 2006, Faux *et al.* 2009, Doherty *et al.* 2009). Yet, despite this around 25% of cases of lung cancer are not being directly linked to smoking (Sun *et al.* 2007). Second hand cigarette smoke, air pollution, inhalation of carcinogens and hereditary genes are all known to increase risk of developing lung cancer (Samet *et al.* 2009). These all also contribute to the multiple respiratory diseases burdening health worldwide (Bousquet, Khaltaev and Cruz, 2007). Chronic obstructive pulmonary disease (COPD), respiratory tract infections, pneumonia and asthma, all contribute to increasing financial cost and strain on medical care with COPD prevalence as high as 9% in some US states (CDC, 2014).

The lack of progress in lung cancer therapeutics combined with international goals for replacement, reduction and refinement (i.e. 3Rs; Russell and Burch, 1959) of animal testing (APC, 2003), is a driving need to shift towards alternative models (BéruBé *et al.* 2011; BéruBé, 2013). These alternative models have historically been limited to monolayer cultures of specific cell lines, which fail to account for the intricacy and real world variability offered through animal testing (Kroll *et al.* 2009). However, a complication associated with *in vivo* testing arises from intra-subject variation of immune responses, which is often activated by foreign object particulates or pathogens (Cressler *et al.* 2014). Use of *in vitro* testing can, as such, simplify understanding of pathological pathways to the route mechanism and is particularly useful in multi-stimuli studies.

Only in recent years have there been advances in the growth of complex tissues capable of providing the intermediary between simple *in vitro* cell monolayers and complex *in vivo* animal models (Bredenkamp *et al.* 2014). Laboratory grown models reduce the burden on animal testing, allowing for multiple cell-type interactions to be examined without the more confounding aspects resulting from the presence of systemic systems and the immune response. Furthermore, the ability to test multiple cell type response using human tissues, negating the reliance on use of alternative species as models, provides more ethically and biologically relevant research into the carcinogenic and damaging effects of inhaled toxicants such as cigarette smoke and air pollution (Adam *et al.* 2015). MatTek's EpiAirway® is a multi-cellular, differentiated model of the human bronchial epithelium derived from healthy human primary tracheo-bronchial cells. The model is aimed at replicating the epithelial tissue of the human respiratory tract (BéruBé *et al.* 2009; Prytherch and BéruBé, 2014).

This study was aimed at confirming, through transcriptomics, the suitability of a laboratory grown human lung model (EpiAirway®) as an alternative model to study the genetic effects on gene regulation and associated pathways caused from tobacco smoke inhalation (Balharry *et al.* 2008; Sexton *et al.* 2008; Sexton *et al.* 2011). The study also aimed to look at the downstream implications on offspring carried during exposure. The EpiAirway® model and pregnant female mice were exposed to mainstream cigarette smoke (MCS) or filtered air before lung tissue RNA extraction and gene regulation were assessed via the Agilent Single Colour Microarray (Agilent, 2015).

Materials and methods

In vitro cell culture exposures

The EpiAirway cell cultures were transported from MatTek in the USA in a 24-well plate format. The cells were equilibrated at 37°C with 4.5% CO_2 for 24 h following manufacturer's guidelines. Culture pre-conditioning, acute exposure (24 h) of the EpiAirway lung tissue (ELT) to mainstream cigarette smoke (MCS) was carried out at the air-liquid interface (ALI) as per the methods by Sexton, Balharry and co-workers (2008), respectively. Each insert had a surface area of 1 cm^2 and was apically dosed.

Animal exposures

Mainstream Cigarette Smoke was generated through burning 3R4F reference filtered cigarettes (Kentucky Tobacco Research and Development Centre, Lexington, KY) on an automated CS generation system (Baumgartner-Jaeger CSM 2070; CH Technologies Inc., Westwood, NJ) and both Dams and EpiAirway® lung tissue cultures were exposed as described in Ng *et al.* (2006). Smoke was drawn from the cigarettes under ISO standard conditions (35 ml puff drawn over 2 seconds every 1 minute) and diluted in filtered air using an RM20s smoke engine (Borgwaldt Technik GmbH, Hamburg, Germany). Diluted smoke (1/50 smoke:air v/v) was continually delivered to exposure chambers (UK patent number WO 03/100417 A1) containing the culture inserts for a period of 1 hr. Exposure was designed to be equivalent to an adult human smoking ~10 cigarettes/day for 18 consecutive days. In the absence of cells, the total deposition of particulates on the base of the cell culture inserts was determined to be 1.84 μg/cm^2. We generated 1 mg/m^3 for the cell culture studies. For our cigarette smoke animal studies, we exposed 4 hr/d running a continuous 15 mg/m^3 concentration in the chamber. $B_6C_6F_1$ mice (Jackson Laboratory, Bar Harbor, ME) were acclimatised and mated as described in Ng *et al.* (2006). Pregnant females (Dams) were exposed to MCS (or filtered air) via whole body inhalation 4 hours/day for 5 days/week during gestation until parturition (18 days) and sacrificed post-exposure and the Dams' and the Offspring' lungs (n = 3 replicates each) were extracted. EpiAirway® lung tissue (ELT; n = 3 replicates) was exposed to the same level of MCS (or filtered air) and for the same duration. Lungs were preserved in RNAlater (Qiagen, USA) for genomic analysis at Cardiff University (Wales).

Sample preparation

Total RNA was extracted and purified from Dam and ELT tissues using RNeasy Mini Kit, (Qiagen, UK). Purity and integrity of extracted RNA was assessed using an Agilent 2100 BioAnalyzer (Agilent Technologies, Palo Alto, CA, USA). Five hundred ng of total RNA from each of the lung tissue samples and the Universal Human/mouse Reference RNA (Stratagene, La Jolla, CA, USA) was used for amplification or RNA and labelling with cyanine (Cy) 5 (experimental samples) or Cy 3 (reference) using the Agilent's Low RNA Input Linear Amplification Kit (Agilent Technologies) according to the manufacturer's instructions. Labelled samples and reference cRNAs were purified using RNeasy mini spin columns (Qiagen, UK) and eluted in 30 μl of nuclease-free water. After amplification and labelling, cRNA quantity and Cy dye incorporation were determined using a Nanodrop ND.1000 UV-VIS-Spectrophotometer version 3.2.1 (Agilent Technologies). For each hybridisation, 1 μg Cy 3 labelled cRNA (reference) and 1 μg of Cy 5 labelled cRNA (samples) were mixed, fragmented, and hybridized at 65°C for 17 hours onto Agilent Whole human/mouse genome 4×44 K 60mer Oligo Microarrays. Labelled cRNA from three different ELT, Dams or Offspring tissues were each hybridized to the arrays. After washing, microarrays were scanned using an Agilent Array scanner (G2505C) (Agilent Technologies) and the images were analysed. Reproducibility and reliability of each single microarray was assessed using Quality Control report data. Data were extracted using Agilent feature extraction software (version 9.5.3) and the GE2-v5_95_Feb07 protocol. Additionally, genes with either uniformly low expression or low expression variation across the experiments were eliminated.

Analysing gene expression

Microarray data was analysed using GeneSpring (version 13.0) to highlight differentially expressed genes. Samples were grouped by exposure to MCS and filtered air for the Dams, Offspring and the ELT, normalising the arrays to the 75th percentile. Quality control on each data set was performed to minimise false detection rate. A moderated T-Test used a cut off p-value value of 0.05 and minimum 1.4 fold change without false detection rate (FDR). These values were chosen as Dalman *et al.* (2011) concluded that lower fold change cut-off produces more significant results in gene ontology (GO). Up and down regulated genes which satisfied these criteria underwent gene enrichment analysis. FDR was accounted for through the use of Gene enrichment of DAVID's Benjamini-Hochberg score for corrected P-

value, rather than through the use of Bonferroni style approaches at earlier stages which, while reducing false positives, often simultaneously exclude true positives. GO-terms and significant pathways for the up and down regulated genes were identified through the use of DAVID (version 6.7) (Huang *et al.* 2009a; Huang *et al.* 2009b) and strength of association assessed though p-value and Benjamini score, where Gene ontology was deconstructed by biological process using REVIGO (Supek *et al.* 2011) with an allowed similarity of 0.7 and visualised using Cytoscape (version 3.2.1) (Shannon *et al.* 2002).

Results

Following normalisation and quality control, 27758, 26571 and 25250 out of 44,000 features were retained for the ELT, Dams and Offspring respectively. Following a moderated T-Test with p-value 0.05 cut-off and minimum 1.4 fold change; 716 (500 up-regulated, 216 down-regulate) genes in the ELT, 437 (283 up-regulated, 154 down-regulated) genes in the Dams and 9825 (5208 up-regulated, 4617 down-regulated) genes in the Offspring were identified as significantly altered, comparing exposure of MCS and filtered air (Figure 1). The top 10 differentially (up and down) expressed genes for the ELT, Dams and Offspring were identified and listed in the Supplemental Table 1.

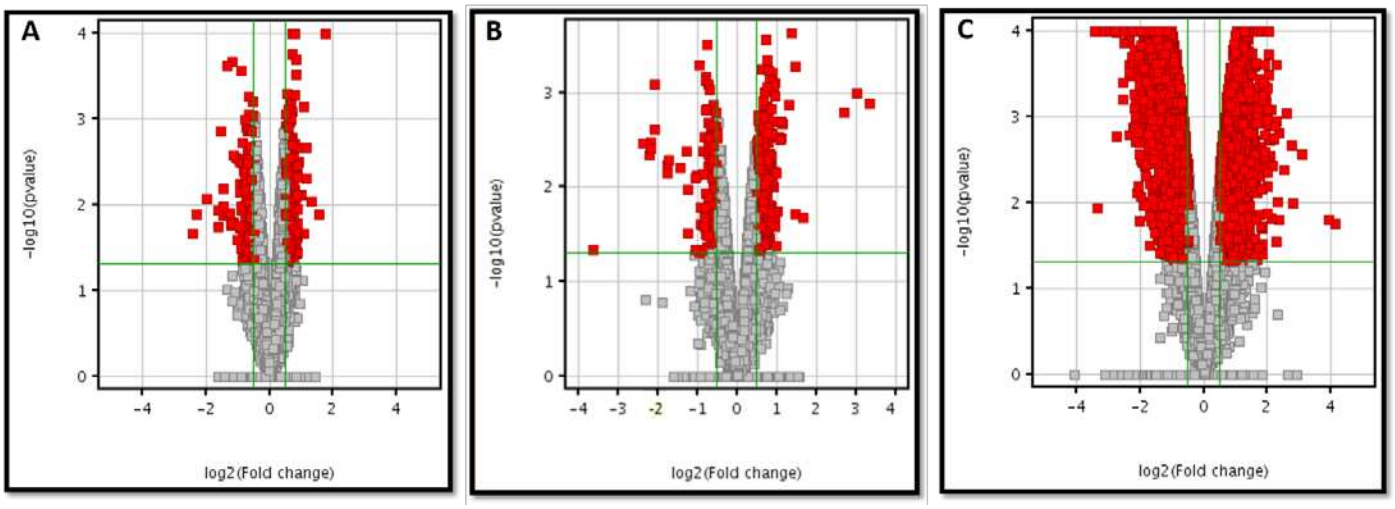


Figure 1. Selection of differentially expressed genes of ELT tissue (A) Dam lung tissue (B) and Offspring lung tissue (C). Volcano plot of differentially expressed genes between ELT, Dams and Offspring tissues exposed to CMS and filtered air. The vertical lines correspond to 1.4-fold up and down expression and the horizontal line represents a p-value of 0.05.

Of the top up-regulated genes in ELT, many have direct functions in regulation and cell adhesion [C-FOS], cell division [CDC20B] and matrix proteins [MATN1], while the top ELT down-regulated genes have links to cell binding [FN1], calcium/zinc ion binding in proteolysis [MMP12] and calcium ion binding in protease inhibition [SPOCK1]. The Dams top up-regulated genes, most have direct functions in immunity, (e.g. Ighg – Immunoglobulin heavy constant-γ), but also includes killer cell lectin-like receptors [KLRA17] and Mediterranean fever [MEFV], with the top Dam down-regulated genes have links to fat regulation and cell life regulation [RETN], regulation of lipid biosynthetic process [THRSP] and mucus production [MUC5B]. The top Offspring up-regulated gene have functions in histocompatibility, [H2AB1], haemoglobin [HBB-BT] and immunity [LY6D]) and the top Offspring down-regulated genes have links to chloride ion channels [BEST1], Tight junctions [TJP2] and GTPases [AGAP1].

The genes that saw the greatest fold change in expression give an indication into large gene network pathways and therefore, gene ontology was performed to see what the global trends in gene regulation of the cell were involved. Up and down regulated genes underwent GO analysis through DAVID and the top 10 enriched terms for the genes associated with, biological process (BP), cellular component (CC) and molecular function (MF) were identified (Table 1).

Table 1. Top 10 Enrichment terms extracted from DAVID.

	Term	P Value		Term	P Value		Term	P Value
ELT BP	Microtubule-Based Movement	8.60E-09	Dams BP	Immune Response	6.70E-19	Offspring BP	Cellular Process	1.23E-91
	Microtubule-Based Process	2.31E-06		Immune System Process	9.00E-16		Cellular Metabolic Process	1.30E-52
	Cellular Developmental Process	7.30E-06		Response To External Stimulus	4.09E-09		Cellular Metabolic Process	1.86E-43
	Ciliary Or Flagellar Motility	1.17E-05		Response To Stimulus	7.60E-09		Metabolic Process	2.31E-41
	Ectoderm Development	1.94E-05		Nuclear Division	9.73E-09		Macromolecule Metabolic Process	1.82E-35
	Developmental Process	2.70E-05		Mitosis	9.73E-09		Primary Metabolic Process	9.05E-35
	Multicellular Organismal Process	5.30E-05		M Phase Of Mitotic Cell Cycle	1.35E-08		Cellular Protein Metabolic Process	8.79E-34
	Cell Differentiation	9.25E-05		Organelle Fission	1.72E-08		Gene Expression	1.64E-25
	Epidermis Development	1.16E-04		Cell Division	3.80E-08		Protein Metabolic Process	5.00E-23
	Tissue Development	2.23E-04		Chemotaxis	7.76E-08		Cellular Component Organization	1.53E-21
ELT CC	Cytoskeleton	1.36E-14	Dams CC	Cell Surface	4.25E-10	Offspring CC	Intracellular	3.90E-135
	Cilium	3.87E-11		External Side Of Plasma Membrane	1.37E-09		Intracellular Part	4.56E-124
	Axoneme	1.47E-10		Extracellular Space	7.40E-08		Cytoplasm	5.02E-96
	Cell Projection	2.06E-10		Chromosome, Centromeric Region	9.11E-07		Organelle	8.86E-83
	Cytoskeletal Part	2.08E-10		Extracellular Region Part	2.18E-06		Intracellular Organelle	1.40E-82
	Microtubule	6.81E-10		Condensed Chromosome, Centromeric Region	1.31E-05		Membrane-Bounded Organelle	7.70E-59
	Microtubule	6.29E-09		Condensed Chromosome Kinetochore	4.88E-05		Intracellular Membrane-Bounded Organelle	1.33E-58
	Cytoskeleton	2.32E-08		Extracellular Region	5.70E-05		Cytoplasmic Part	1.25E-52
	Microtubule Associated Complex	6.26E-08		Chromosomal Part	8.49E-05		Intracellular Organelle Part	8.85E-38
	Cell Projection Part	8.15E-08		Kinetochore	1.35E-04		Organelle Part	7.09E-37
	Dynein Complex							
ELT MF	Microtubule Motor Activity	4.25E-09	Dams MF	Sugar Binding	2.08E-11	Offspring MF	Binding	1.21E-65
	Structural Molecule Activity	2.94E-07		Carbohydrate Binding	2.15E-10		Protein Binding	8.57E-42
	Motor Activity	3.72E-07		Chemokine Activity	5.32E-06		Structural Constituent Of Ribosome	2.72E-26
	Metalloendopeptidase Activity	4.56E-05		Chemokine Receptor Binding	6.39E-06		Cytoskeletal Protein Binding	2.27E-16
	Metallopeptidase Activity	1.60E-04		Cytokine Activity	6.58E-06		Catalytic Activity	1.04E-14
	Calcium Ion Binding	2.67E-04		Receptor Binding	6.38E-05		Purine Ribonucleotide Binding	3.90E-14
	Structural Constituent Of Cytoskeleton	4.27E-04		Sh3/Sh2 Adaptor Activity	2.26E-04		Ribonucleotide Binding	3.90E-14
	Nucleoside-Triphosphatase Activity	0.003537		Signal Transducer Activity	2.34E-04		Purine Nucleotide Binding	1.13E-13
	Extracellular Matrix Structural Constituent	0.004979		Molecular Transducer Activity	2.34E-04		Actin Binding	1.25E-13
	Pyrophosphatase Activity	0.006094		Protein Binding	2.89E-04		Nucleotide Binding	2.24E-13

DAVID ranks associated genes with a GO term to provide an enrichment score, with highly enriched terms have a greater number of associated genes. Many of the top 10 BPs, MFs and all the CCs for ELT relate to cytoskeletal genes and cell development. The Dams had BPs relating to cell cycle and additionally significant immune response alterations. The CCs identified were mostly involved with the chromosomal organisation while MFs had highly enriched terms in receptor binding and transmission. This suggested that ELT was primarily affected in cellular organisation while the Dams were primarily responding to external stimuli, above a cellular organisation response. To investigate similarities in response, common GO terms between ELT and the Dams were collated. Offspring had many GO terms associated with metabolic processes, organelle structure and protein binding.

To assess the similarities between ELT as a model for replacement of the Dams, common processes for all GO-terms for ELT and Dams were identified (Table 2). The common terms can largely be linked to the processes of cellular adhesion and response to a stimulus. This indicated the Dams immune response was masking similar mechanical pathway changes with the ELT. The global KEGG pathways identified through DAVID display the complexity of the Dams when compared to the ELT (Table 3). ELT tissues highlight 5 pathways, including ECM-receptor interactions and pathways in cancer and cardiac stress. Dam tissue had a large range of 19 disease pathways altered that included immunity, cell cycle regulation, diabetes and cell adhesion. The Offspring however showed a vast network of 61 disease and cancer pathways which included Alzheimer's and Parkinson's, Chronic myeloid leukaemia, Colorectal cancer and Type II diabetes.

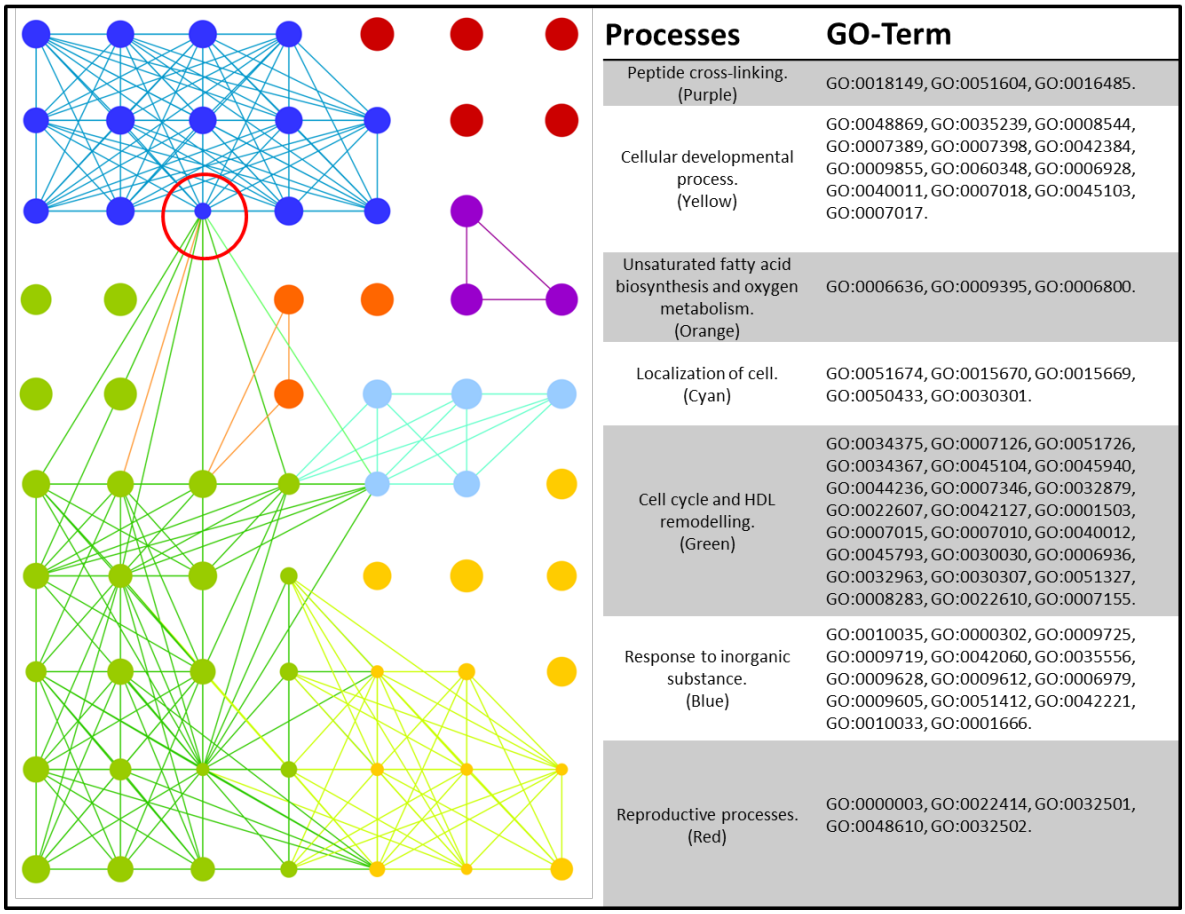
Table 2. Common GO terms found between ELT and Dams.

GO term	Description	ELT P-value	Dams P-value
GO:0044421	Extracellular region part	3.41E-06	2.18E-06
GO:0005576	Extracellular region	8.30E-05	5.70E-05
GO:0005615	Extracellular space	6.59E-04	7.40E-08
GO:0005488	Binding	2.16E-02	1.70E-03
GO:0044459	Plasma membrane part	2.87E-02	9.75E-03
GO:0016043	Cellular component organization	3.42E-02	5.07E-03
GO:0042060	Wound healing	3.68E-02	5.18E-02
GO:0042221	Response to chemical stimulus	3.89E-02	3.02E-02
GO:0009605	Response to external stimulus	4.10E-02	4.09E-09
GO:0007155	Cell adhesion	4.10E-02	4.67E-02
GO:0022610	Biological adhesion	4.17E-02	4.77E-02
GO:0016485	Protein processing	4.88E-02	7.93E-02
GO:0040011	Locomotion	4.88E-02	9.39E-04
GO:0042127	Regulation of cell proliferation	5.01E-02	6.26E-02
GO:0032879	Regulation of localization	5.34E-02	4.75E-05
GO:0051604	Protein maturation	6.81E-02	9.50E-02
GO:0030674	Protein binding, bridging	7.96E-02	3.08E-03
GO:0045834	Positive regulation of lipid metabolic process	9.35E-02	6.88E-02

Table 3. Differentially regulated gene interaction identified in Kegg pathways

	Kegg term	Gene count	-LogP	Kegg term	Gene count	-LogP
ELT	Focal adhesion	12	2.78	ARVC	5	1.22
	ECM-receptor interaction	8	2.94	Small cell lung cancer	5	1.10
	p53 signaling pathway	6	1.99	Oocyte meiosis	5	1.07
	Calcium signaling pathway	8	1.26			
Dams	Natural killer cell mediated cytotoxicity	20	11.81	Systemic lupus erythematosus	7	1.79
	Graft-versus-host disease	13	9.06	ECM-receptor interaction	6	1.62
	Cytokine-cytokine receptor interaction	21	7.33	NOD-like receptor signaling pathway	5	1.47
	Hematopoietic cell lineage	9	3.63	Antigen processing and presentation	6	1.47
	Chemokine signaling pathway	12	3.04	p53 signaling pathway	5	1.32
	Allograft rejection	7	3.03	Autoimmune thyroid disease	5	1.27
	Type I diabetes mellitus	7	2.84	Oocyte meiosis	6	1.11
	Circadian rhythm	4	2.72	T cell receptor signaling pathway	6	1.07
	Toll-like receptor signaling pathway	8	2.47	Cell adhesion molecules (CAMs)	7	1.07
	Cytosolic DNA-sensing pathway	6	2.35			
Offspring	Focal adhesion	110	11.70	Purine metabolism	64	1.94
	Huntington's disease	99	9.65	Nucleotide excision repair	22	1.91
	ECM-receptor interaction	49	6.31	Glioma	30	1.89
	Oxidative phosphorylation	68	5.96	Non-small cell lung cancer	26	1.83
	Adherens junction	44	5.37	RNA polymerase	15	1.68
	Alzheimer's disease	86	5.09	Colorectal cancer	37	1.60
	Ubiquitin mediated proteolysis	68	5.08	Propanoate metabolism	16	1.60
	Regulation of actin cytoskeleton	99	5.00	SNARE interactions in vesicular transport	19	1.55
	Axon guidance	64	4.40	mTOR signaling pathway	25	1.54
	Endocytosis	91	4.36	Type II diabetes mellitus	23	1.50
	Parkinson's disease	64	4.15	Leukocyte transendothelial migration	48	1.45
	Proteasome	28	3.73	B cell receptor signalling pathway	34	1.41
	Viral myocarditis	47	3.59	Graft-versus-host disease	26	1.41
	Valine, leucine and isoleucine degradation	27	3.46	p53 signalling pathway	30	1.41
	Small cell lung cancer	43	3.44	Long-term potentiation	30	1.32
	Fc gamma R-mediated phagocytosis	48	3.40	Hypertrophic cardiomyopathy (HCM)	35	1.32
	Spliceosome	58	3.39	ErbB signalling pathway	36	1.31
	Insulin signalling pathway	62	3.03	Cell adhesion molecules (CAMs)	59	1.26
	Pancreatic cancer	36	2.80	Glycosylphosphatidylinositol(GPI)	13	1.22
	Renal cell carcinoma	35	2.73	-anchor biosynthesis		
	Pyrimidine metabolism	45	2.71	Lysine degradation	19	1.21
	Vascular smooth muscle contraction	53	2.45	Gap junction	35	1.18
	Glutathione metabolism	27	2.42	Dilated cardiomyopathy	37	1.16
	RNA degradation	30	2.37	(ARVC)	31	1.16
	Tight junction	58	2.33	Aldosterone-regulated sodium reabsorption	19	1.11
	Pathways in cancer	124	2.30	Natural killer cell mediated cytotoxicity	47	1.10
	Prostate cancer	41	2.23	Cell cycle	49	1.09
	Phosphatidylinositol signalling system	35	2.13	Fatty acid biosynthesis	5	1.09
	Neurotrophin signalling pathway	55	2.06	GnRH signalling pathway	38	1.04
	Chronic myeloid leukemia	35	2.02	Systemic lupus erythematosus	40	1.03
	Apoptosis	39	2.00			

The links between BP, CC and MF GO-terms processes were assessed utilising REVIGO and visualised in Cytoscape. ELT BPs showed several clusters of GO-terms (Figure 2). The large group GO-terms have been mapped by colour and listed with their associated GO-terms. One group classified under HDL particle remodelling included many cellular organisation processes and were heavily interlinked with cellular development processes. Response to inorganic substance formed many intragroup links but connected to HDL particle remodelling through a single node process of intracellular signal transduction. The majority of ELT CCs are interlinked cytoskeletal processes and extracellular structures and basolateral plasma membrane while MFs had interlinks between molecular binding, structural activity and kinase/transferase activity (Supplemental Table 2).



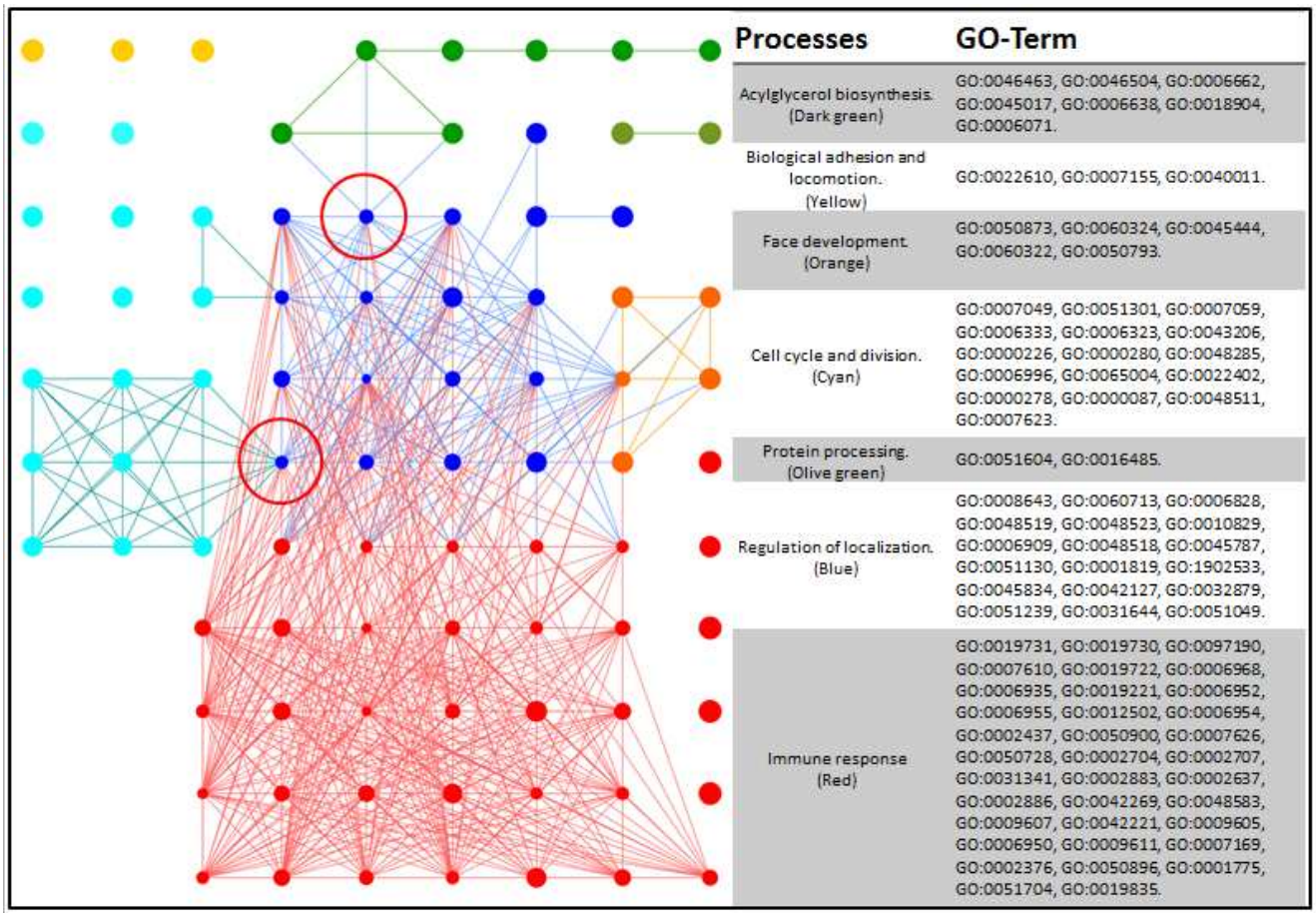
Processes	GO-Term
Peptide cross-linking. (Purple)	GO:0018149, GO:0051604, GO:0016485.
Cellular developmental process. (Yellow)	GO:0048869, GO:0035239, GO:0008544, GO:0007389, GO:0007398, GO:0042384, GO:0009855, GO:0060348, GO:0006928, GO:0040011, GO:0007018, GO:0045103, GO:0007017.
Unsaturated fatty acid biosynthesis and oxygen metabolism. (Orange)	GO:0006636, GO:0009395, GO:0006800.
Localization of cell. (Cyan)	GO:0051674, GO:0015670, GO:0015669, GO:0050433, GO:0030301.
Cell cycle and HDL remodelling. (Green)	GO:0034375, GO:0007126, GO:0051726, GO:0034367, GO:0045104, GO:0045940, GO:0044236, GO:0007346, GO:0032879, GO:0022607, GO:0042127, GO:0001503, GO:0007015, GO:0007010, GO:0040012, GO:0045793, GO:0030030, GO:0006936, GO:0032963, GO:0030307, GO:0051327, GO:0008283, GO:0022610, GO:0007155.
Response to inorganic substance. (Blue)	GO:0010035, GO:0000302, GO:0009725, GO:0009719, GO:0042060, GO:0035556, GO:0009628, GO:0009612, GO:0006979, GO:0009605, GO:0051412, GO:0042221, GO:0010033, GO:0001666.
Reproductive processes. (Red)	GO:0000003, GO:0022414, GO:0032501, GO:0048610, GO:0032502.

Figure 2. ELT – BP interaction map. GO terms processed through REVIGO are visualised through Cytoscape. Node sizes are correlated to the 'uniqueness' value determined by REVIGO, where smaller nodes share more similarity with neighbouring GO terms. The red circle indicates bottleneck between the 'Response to inorganic substance's cluster and the larger and more heavily interlinked 'Cell cycle and HDL remodelling'.

The Dams BP associated REVIGO groupings showed a far more complex mapping (Figure 3). The large group GO-terms have been mapped by colour and listed with their associated GO-terms. There is a vast and heavily interconnected cluster of immune processes which also form multiple interactions with regulation of localisation. Smaller clusters of nuclear division and acylglycerol biosynthesis connect through to this massive cluster through single nodes of positive regulation of cellular component organisation and of lipid metabolism. There was a small CC interconnection with processes relating to the cell surface, extracellular space and protein/DNA interaction. MF interconnections between cytokine activity, Protein activity and Binding was also identified (Supplemental Table 3).

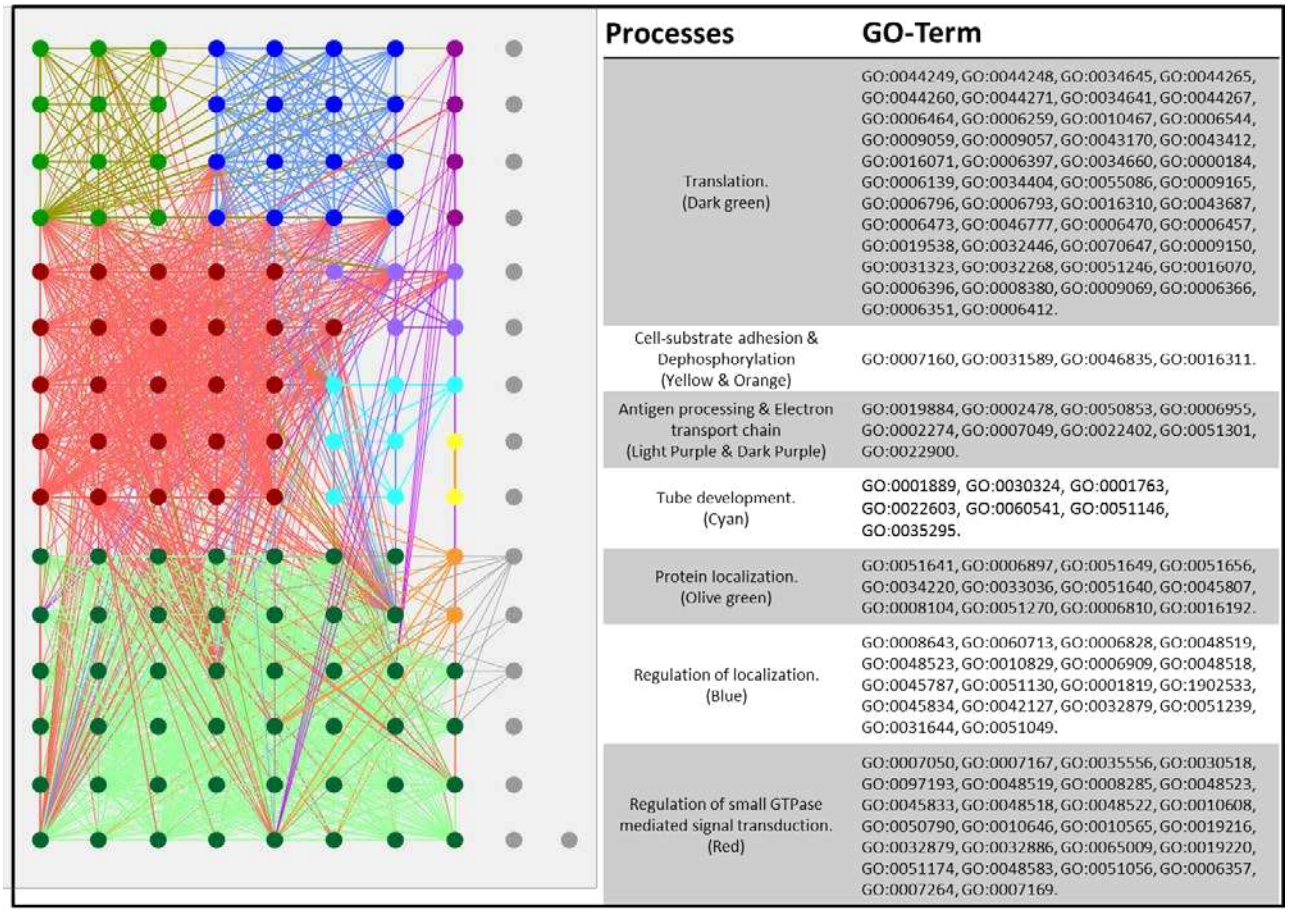
Large similarities between the ELT and Dams exist in Cell cycle regulation and localization. While the ELT sees alteration to processes associated with response to inorganic substances on a cellular level, the Dams have a heavy immune response as a whole. The influence of this immune response can distract from the cellular mechanistic responses and is outlined in greater detail in the discussion.

The Offspring REVIGO map of associated GO terms displayed a vast and highly interconnected network both intra-processes and inter-processes (Figure 4). The large group GO-terms have been mapped by colour and listed with their associated GO-terms. Translation and regulation of GTPases had the largest networks of GO terms in the BP map, while there was also a large number of regulatory changes in localization, cell-substrate adhesion and tube development and antigen processing. There was a large network of GO terms associated with the Mitochondrion cell projection and the basement membrane for the CC map and cytoskeletal binding, motor activity, ubiquitin-protein transferase activity and zinc iron binding in the MF map (Supplemental Table 4).



Processes	GO-Term
Acylglycerol biosynthesis. (Dark green)	GO:0046463, GO:0046504, GO:0006662, GO:0045017, GO:0006638, GO:0018904, GO:0006071.
Biological adhesion and locomotion. (Yellow)	GO:0022610, GO:0007155, GO:0040011.
Face development. (Orange)	GO:0050873, GO:0060324, GO:0045444, GO:0060322, GO:0050793.
Cell cycle and division. (Cyan)	GO:0007049, GO:0051301, GO:0007059, GO:0006333, GO:0006323, GO:0043206, GO:0000226, GO:0000280, GO:0048285, GO:0006996, GO:0065004, GO:0022402, GO:0000278, GO:0000087, GO:0048511, GO:0007623.
Protein processing. (Olive green)	GO:0051604, GO:0016485.
Regulation of localization. (Blue)	GO:0008643, GO:0060713, GO:0006828, GO:0048519, GO:0048523, GO:0010829, GO:0006909, GO:0048518, GO:0045787, GO:0051130, GO:0001819, GO:1902533, GO:0045834, GO:0042127, GO:0032879, GO:0051239, GO:0031644, GO:0051049.
Immune response (Red)	GO:0019731, GO:0019730, GO:0097190, GO:0007610, GO:0019722, GO:0006968, GO:0006935, GO:0019221, GO:0006952, GO:0006955, GO:0012502, GO:0006954, GO:0002437, GO:0050900, GO:0007626, GO:0050728, GO:0002704, GO:0002707, GO:0031341, GO:0002883, GO:0002637, GO:0002886, GO:0042269, GO:0048583, GO:0009607, GO:0042221, GO:0009605, GO:0006950, GO:0009611, GO:0007169, GO:0002376, GO:0050896, GO:0001775, GO:0051704, GO:0019835.

Figure 3. Dams – BP interaction map. GO terms processed through REVIGO are visualised through Cytoscape. Node sizes are correlated to the 'uniqueness' value determined by REVIGO, where smaller nodes share more similarity with neighbouring GO terms. The red circle indicates bottleneck between the 'Cell cycle and division' cluster and the larger and more heavily interlinked 'Immune system'. A bottle neck also connects 'Acylglycerol biosynthesis' pathways with 'Regulation of localisation'.



Processes	GO-Term
Translation. (Dark green)	GO:0044249, GO:0044248, GO:0034645, GO:0044265, GO:0044260, GO:0044271, GO:0034641, GO:0044267, GO:0006464, GO:0006259, GO:0010467, GO:0006544, GO:0009059, GO:0009057, GO:0043170, GO:0043412, GO:0016071, GO:0006397, GO:0034660, GO:0000184, GO:0006139, GO:0034404, GO:0055086, GO:0009165, GO:0006796, GO:0006793, GO:0016310, GO:0043687, GO:0006473, GO:0046777, GO:0006470, GO:0006457, GO:0019538, GO:0032446, GO:0070647, GO:0009150, GO:0031323, GO:0032268, GO:0051246, GO:0016070, GO:0006396, GO:0008380, GO:0009069, GO:0006366, GO:0006351, GO:0006412.
Cell-substrate adhesion & Dephosphorylation (Yellow & Orange)	GO:0007160, GO:0031589, GO:0046835, GO:0016311.
Antigen processing & Electron transport chain (Light Purple & Dark Purple)	GO:0019884, GO:0002478, GO:0050853, GO:0006955, GO:0002274, GO:0007049, GO:0022402, GO:0051301, GO:0022900.
Tube development. (Cyan)	GO:0001889, GO:0030324, GO:0001763, GO:0022603, GO:0060541, GO:0051146, GO:0035295.
Protein localization. (Olive green)	GO:0051641, GO:0006897, GO:0051649, GO:0051656, GO:0034220, GO:0033036, GO:0051640, GO:0045807, GO:0008104, GO:0051270, GO:0006810, GO:0016192.
Regulation of localization. (Blue)	GO:0008643, GO:0060713, GO:0006828, GO:0048519, GO:0048523, GO:0010829, GO:0006909, GO:0048518, GO:0045787, GO:0051130, GO:0001819, GO:1902533, GO:0045834, GO:0042127, GO:0032879, GO:0051239, GO:0031644, GO:0051049.
Regulation of small GTPase mediated signal transduction. (Red)	GO:0007050, GO:0007167, GO:0035556, GO:0030518, GO:0097193, GO:0048519, GO:0008285, GO:0048523, GO:0045833, GO:0048518, GO:0048522, GO:0010608, GO:0050790, GO:0010646, GO:0010565, GO:0019216, GO:0032879, GO:0032886, GO:0065009, GO:0019220, GO:0051174, GO:0048583, GO:0051056, GO:0006357, GO:0007264, GO:0007169.

Figure 4. Offspring – BP interaction map. GO terms processed through REVIGO are visualised through Cytoscape. Node sizes are correlated to the 'uniqueness' value determined by REVIGO, where smaller nodes share more similarity with neighbouring GO terms. The highly interconnecting map does not have any bottlenecks in process interaction as seen in the ELT and Dams maps. The nodes in Grey show multiple smaller unlinked GO terms.

Discussion

Despite the high incidence and mortality that accompanies lung cancer and other pulmonary diseases, research funding remains disproportionately low. The additional pressures for refinement, reduction and replacement of animal testing is a driving need for alternative models that can more accurately represent human *in vivo* responses. This pilot study aimed at determining if the ELT (EpiAirway®) could provide this alternative without the obfuscating presence of an immune system. It was hoped the ELT would provide a reductionist view, useful for understanding initial, site-of impact and mechanics of the cellular interactions.

Global transcriptomics pathway analysis allows analysis between individuals and also cross-species comparison. Requiring higher fold changes has often been used to filter out normal fluctuations in gene regulation. This accompanied with early FDR compensation can exclude many relevant genes. Indeed, some genes only require a minimal fluctuation in their regulation to have a profound effect and are highly regulated to avoid fluctuations (Raser and O'shea, 2005). Filtering genes with a lower threshold for fold change and allowing significance to be assessed in gene enrichment processes provides a more informative and reliable picture of global cellular response (Dalman *et al.* 2011).

Comparatively, fewer differentially expressed genes met the criteria for gene enrichment in the ELT model than in the Dams. The top differentially regulated genes and their functions were initially identified. In the ELT many of the upregulated genes had functions linked to cell cycle, while there was downregulation of cell adhesion and calcium

homeostasis regulation. In the Dams there was heavy upregulation of immunoglobulins and immunity receptors and down regulation of fat regulation, detection and signalling pathways.

These differences were observed (Table 1) for the 'immune' versus the 'mechanical' damage in the top 10 enriched terms, where the top 10 in the Dams included immune response and response to external stimulus but the ELT is cellular structure. However, there were similarities, suggesting the same mechanical issues are occurring but they are hidden behind the overwhelming immune response. For example, common GO Terms included extra-cellular components (i.e. regions and spaces) and response to chemicals and external stimuli (Table 2). We also observed common Kegg pathways, such as the P53 signalling and ECM interaction pathways (Table 3). More pathways were exacerbated in Dams, again linked to the exceptional immune response, but cell cycle and cell adhesion responses were common. The inter-linking of these pathways demonstrated the inter-connectivity of stress responses and the comparative size of the immune response, but also common stress responses. Histologically, these broad similarities were observed, such as the loss of tight junctions, cytoskeleton differences and inflammatory responses associated with the exacerbated immune response.

With regard to the additional analyses of the Offspring gene changes, though this research focus was the ELT versus Dam model comparison, the heavy alterations (i.e. 9,825 differentially expressed genes), nonetheless provides an interesting data-set. For example, GO Terms included those associated largely with translation, regulation of small GTPase-mediated signal transduction and regulation of localisation (Figure 4). Foetal development is a time when many genes are being turned on and off and the impacts of chemical exposures might well explain the significant number that was observed. Many of the known later-life impacts following pre-natal exposures to CS have been identified (Table 3), such as diabetes, Alzheimer's and multiple cancer pathways (Doherty *et al.* 2009).

Often times, animal models don't recapitulate what is observed epidemiologically when it comes to cigarette smoke-induced carcinogenesis. However, a study by Hutt et al (2005; Carcinogenesis 26:11) used the $B_6C_3F_1$ mouse strain (the same used in the present study) and found that lifetime exposure of female mice to MCS to 250 mg PM/m^3 for 6 h/d, 5 d/wk, induced an increased rate of focal alveolar hyperplasia, pulmonary adenomas, papillomas and adenocarcinomas. Versus unexposed control mice, those exposed to MCS had 10-fold increase in hyperplastic lesions, 4.6-fold increase in adenomas, 7.25-fold increase in adenocarcinomas, and 5-fold increase in metastatic pulmonary adenocarcinoma. The selection of the mouse strain, i.e. the $B_6C_3F_1$ hybrid strain, has been used in carcinogenesis assays by many researchers, as well as the National Toxicology Program (USA) due to its lung tumour response to certain chemicals like cigarette smoke and chemicals present in cigarette smoke. In addition, as reviewed by Pandiri (2015), "Meta-analysis of transcriptomic alterations in human and mouse lung tumours revealed significant similarities in lung cancer pathways in both species (Stearman *et al.* 2005; Bonner *et al.* 2004; Pandiri *et al.* 2012). These data indicate that mouse lung tumours are similar to human adenocarcinomas at the morphologic and molecular levels and that mouse lung tumours are relevant in evaluating carcinogenic hazards associated with environmental exposures."

In conclusion, the gene changes observed in the *in vitro*, 3-dimensional, cell culture model of the human bronchial epithelium mirrored the responses detected in the mouse model of pre-natal exposure to MCS. The ELT model could be utilised as the first step (i.e. before using animal models) to screening aerosolised compounds such as candidate respiratory drugs (Prytherch *et al.* 2011), combustion-derived air pollution (e.g. tobacco smoke (Sexton *et al.* 2008; Balharry *et al.* 2008; Sexton *et al.* 2011), diesel exhaust, fly ash particles and shipping emissions (Oder *et al.* 2015). The benefits of using alternative *in vivo*-like *in vitro* ALI models of the human lung are self-evident. The ELT model is both cost- and time-effective for toxicity testing of aerosolised and soluble compounds given that cell culture consumables are highly-affordable and permit rapid analyses. In comparison to animal models that are expensive due to costs of the animals and their maintenance, which could last for years, versus days and/or weeks for *in vitro* cell culturing practices (Prytherch and BéruBé, 2014). Finally, when considering the contentious ethical issues surrounding the use of animals for medical research, when using alternative systems like MatTek's EpiAirway® platform, there is immediate impact for the 3Rs and human end-point data is acquired, negating the need to extrapolate data from animals into effects in man.

References

1. Agilent (2015). One-Color Microarray-Based Gene Expression Analysis Low Input Quick Amp Labeling Protocol Version 6.9.1. http://www.agilent.com/cs/library/usermanuals/Public/G4140-90040_GeneExpression_OneColor_6.9.pdf
2. Adam, M., Schikowski, T., Carsin, A. E., Cai, Y., Jacquermin, B., Sanchez, M., … Probst-Hensch, N. (2015) Adult lung function and long-term air pollution exposure. ESCAPE: a multicentre cohort study and meta-analysis. *European respiratory journal*. **45**(1): 38-50.
3. Animal Procedures Committee APC (2003) *Review of cost-benefit assessment in the use of animals in research June 2003*. Home Office publication, Communication Directorate.
4. Balharry, D., Sexton, K. J. and BéruBé, K. A. 2008. An *in vitro* approach to assess the toxicity of inhaled tobacco smoke components: nicotine, cadmium, formaldehyde and urethane. *Toxicology* 244(1), pp. 66-76. (10.1016/j.tox.2007.11.001)
5. BéruBé, K., Prytherch, Z., Job, C., Hughes, T. (2009) Human primary bronchial lung cell constructs: The new respiratory models. *Toxicology*. **278**(3): 311-318.
6. BéruBé, K., Pitt, Aldo, Hayden, Patrick, Prytherch, Zoë Cariad and Job, Claire Alison (2010). Filter-well technology for advanced three-dimensional cell culture: perspectives for respiratory research. *Alternatives to Laboratory Animals* 38(S1), pp. 49-65.
7. BéruBé, K., Gibson, C., Job C. and Prytherch, Z. (2011). Human lung tissue engineering: a critical tool for safer medicines. *Cell and Tissue Banking* 12(1), pp. 11-13. (10.1007/s10561-010-9204-6)
8. BéruBé, K. A. 2013. Medical waste tissues - breathing life back into respiratory research. *ATLA: Alternatives to Laboratory Animals* 41(6), pp. 429-434.
9. Bonner, A. E., Lemon, W. J., Devereux, T. R., Lubet, R. A., and You, M. (2004). Molecular profiling of mouse lung tumors: Association with tumor progression, lung development, and human lung adenocarcinomas. Oncogene 23, 1166–76.
10. Bredenkamp, N., Ulyanchenko, S., O'Neill, K. E., Manley, N. R., Vaidya, H. J., and Blackburn, C. C. (2014). An organized and functional thymus generated from FOXN1-reprogrammed fibroblasts. *Nature cell biology.* **16:** 902-908.
11. Bousquet, J., Khaltaev, N. and Cruz, A. (2007) Global surveillance, prevention and control of chronic respiratory diseases. Geneva: World Health Organization.
12. Centre for disease control and prevention, U. S. (CDC) (2014) Chronic obstructive pulmonary disease (COPD) – Data and Statistics. *Available at http://www.cdc.gov/copd/data.htm.* (Accessed 4th August 2015).
13. Cressler, C. E., Nelson, W. A., Day, T., McCauley, E., and Bonsall, M. (2014) Disentangling the interaction among host resources, the immune system and pathogens. *Ecology letters*. **17**(3): 284-293.
14. Dalman, M. R., Deeter, A., Nimishakavi, G. and Duan, A. (2011) Fold change and p-value cutoffs significantly alter microarray interpretations. *BMC Bioinformatics*. **13**(2): S11.
15. Doherty, S. P., Grabowski, J., Hoffman, C., Ng, S. P., and Zelikoff, J.T. (2009) Early life insult from cigarette smoke may be predictive of chronic diseases later in life. *Biomarkers*. **14**(S1): 97-101.
16. Faux, S. P., Tai, T., Throne, D., Xu, Y., Breheny, D., and Gaca, M. (2009) The role of oxidative stress in the biological responses of lung epithelial cells to cigarette smoke. *Biomakers.* **14**(S1): 90-96.
17. Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009a) Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protocols*. **4**(1):44-57.
18. Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009b) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*. **37**(1):1-13.
19. Hutt, J. A., Vuillemenot, B. R., Barr, E. B., Grimes, M. J., Hahn, F. F., Hobbs, C. H., March, T. H., Gigliotti, A. P., Seilkop, S. K., Finch, G. L., Mauderly, J. L., and Belinsky, S. A. (2005). Life-span inhalation exposure to mainstream cigarette smoke induces lung cancer in B6C3F1 mice through genetic and epigenetic pathways. Carcinogenesis 26, 1999–2009.

20. Kroll, A., Pillukat, M. H., Hahn, D., and Schnekenburger, J. (2009) Current *in vitro* methods in nanoparticle risk assessment: Limitations and challenges. *European Journal of Pharmaceutics and Biopharmaceutics.* **72**(2): 370-377.
21. National Institutes of Health, U. S. (NIH) (2015) Estimates of funding for various research condition and disease categories (RCDC) NIH research portfolio online reporting tools. *Available at http://report.nih.gov/categorical_spending.aspx*. (Accessed: 23rd April 2015).
22. Ng, S. P., Silverstone, A. E., Lai, Z. and Zelikoff, J. T. (2006) Effects of prenatal exposure to cigarette smoke on offspring tumour susceptibility and associated immune mechanisms. *Toxicological sciences*. **89**(1): 135-144.
23. Oeder, S.et al. 2015. Particulate matter from both heavy fuel oil and diesel fuel shipping emissions show strong biological effects on human lung cells at realistic and comparable in vitro exposure conditions. *PLoS ONE* 10(6), article number: e0126536. (10.1371/journal.pone.0126536)
24. Pandiri A. (2015). Comparative pathobiology of environmentally induced lung cancers in humans and rodents. Toxicol Pathol 43(1):107-14.
25. Pandiri, A. R., Sills, R. C., Ziglioli, V., Ton, T. V., Hong, H. H., Lahousse, S. A., Gerrish, K. E., Auerbach, S. S., Shockley, K. R., Bushel, P. R., Peddada, S. D., and Hoenerhoff, M. J. (2012). Differential transcriptomic analysis of spontaneous lung tumors in B6C3F1 mice: Comparison to human non-small cell lung cancer. Toxicol Pathol 40, 1141–59.
26. Prytherch, Z.et al. 2011. Tissue-specific stem cell differentiation in an in vitro airway model. *Macromolecular Bioscience* 11(11), pp. 1467-1477. (10.1002/mabi.201100181)
27. Prytherch, Z. C. and BéruBé, K. A. 2014. A normal and biotransforming model of the human bronchial epithelium for the toxicity testing of aerosols and solubilised substances. *Alternatives to Laboratory Animals: ATLA* 42(6), pp. 377-381.
28. Raser, J. M. and O'shea, E. K. (2005) Noise in gene expression: Origins, consequences and control. *Science*. **39**(5743): 2013-2013.
29. Russell, W.M.S. and Burch, R.L., (1959). *The Principles of Humane Experimental Technique*, Methuen, London. ISBN 0900767782.
30. Samet, J. M., Avila-Tang, E., Boffetta, P., Hannan, L. M, Olivo-Marston, S., Thun, M. J. and Rudin, C. M. (2009) Lung cancer in never smokers: Clinical epidemiology and environmental risk factors. *Clinical cancer research*. **15**(18): 5626-5645.
31. Sexton, K., Balharry, D. and BéruBé, K. A. (2008). Genomic biomarkers of pulmonary exposure to tobacco smoke components. *Pharmacogenetics and Genomics* 18(10), pp. 853-860. (10.1097/FPC.0b013e328307bddf)
32. Sexton, K., Balharry, D., Brennan, P., Brewis, I., and BéruBé, K. (2011). Proteomic profiling of human respiratory epithelia by iTRAQ reveals biomarkers of exposure and harm by tobacco smoke components. *Biomarkers* 16(7), pp. 567-576. (10.3109/1354750X.2011.608855)
33. Siegel, R., Ma, J., Zou, Z., and Jemal, A. (2014) Cancer statistics, 2014. *A Cancer Journal for Clinician.* **64**: 9–29.
34. Stearman, R. S., Dwyer-Nield, L., Zerbe, L., Blaine, S. A., Chan, Z., Bunn, P. A., Jr., Johnson, G. L., Hirsch, F. R., Merrick, D. T., Franklin, W. A., Baron, A. E., Keith, R. L., Nemenoff, R. A., Malkinson, A. M., and Geraci, M. W. (2005). Analysis of orthologous gene expression between human pulmonary adenocarcinoma and a carcinogen-induced murine model. Am J Pathol 167, 1763–75.
35. Supek, F., Bosnjak, M., Skunca, N. and Smuc, T. (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*. **6**(7): e21800.
36. Sun, S., Schiller, J. H., Gazdar, A. F. (2007) Lung cancer in never smokers – a different disease. *Nature reviews cancer*. **7**: 778-790.

Supplemental Information

Supplementary Table 1. Top 10 up and down differentially regulated genes in ELT, Dams and Offspring tissue after MCS exposure.

	Up-regulated genes			Down-regulated genes		
ELT	Associated Gene ID	P-value	Fold change	Associated Gene ID	P-value	Fold change
	C-FOS	8.85E-05	3.3303	FN1	2.16E-02	5.3997
	CD300A	1.29E-02	2.8970	SPARC	8.55E-03	3.9802
	SAG	9.15E-03	2.5003	CRNN	1.18E-02	3.0898
	MATN1	2.19E-03	2.2019	ADAM19	1.38E-03	2.9293
	CXCR4	4.92E-03	2.1926	GPC6	1.06E-02	2.7941
	SCG3	7.13E-04	2.1032	MMP12	6.47E-03	2.7559
	FRMD1	2.14E-02	2.0700	MMP2	1.32E-02	2.7509
	ELMOD1	3.15E-03	1.8967	SRPX	2.36E-04	2.5735
	HERC2P9	2.20E-03	1.8487	APOE	1.33E-02	2.5694
	GLYATL2	3.67E-02	1.8263	SPOCK1	1.24E-02	2.4189
Dams	IGHG	1.28E-03	10.0077	IGKV8	4.55E-02	12.4052
	IGKV5	5.29E-04	2.7783	RETN	4.47E-03	4.5923
	GZMB	2.29E-04	2.5938	TMEM45B	3.93E-03	4.5373
	IGHV1	1.35E-03	2.4647	CAR3	8.08E-04	4.2618
	MTFR2	3.14E-03	2.1057	ADIPOQ	2.43E-03	4.2096
	IL1B	2.90E-03	2.0812	CFD	7.08E-03	3.3732
	HIST1HiB	4.13E-02	1.9907	TFF2	6.26E-03	2.7266
	KLRA17	6.94E-03	1.9576	THRSP	4.08E-03	2.4456
	ARNTL	1.08E-03	1.9509	PLIN1	1.07E-02	2.3777
	FEFV	7.98E-04	1.9452	MUC5B	3.06E-02	2.3440
Offspring	H2AB1	0.017866	17.57977	BEST1	8.17E-07	10.84912
	H2Q7	0.016136	15.42327	MBNL1	0.011763	10.23817
	TRMT61	0.0028	8.509253	AHDC1	8.64E-06	7.887322
	LY6D	0.010242	6.925174	BAIAP2L1	2.17E-05	7.718487
	STRA8	0.002144	6.818615	KCNQ1OT1	9.31E-06	7.359658
	H2DMB2	9.13E-04	6.113622	TNS1	5.16E-06	7.111162
	APOL7C	0.015917	5.175049	TJPs	2.51E-06	6.984259
	IFI44L	0.001521	5.093131	AGAP1	7.56E-06	6.975762
	HBB-BT	0.028783	4.819223	PITPNC1	0.00172	6.867369
	IKZF4	2.50E-04	4.78618	ZFHX3	7.63E-06	6.528332

Supplemental Table 2. Some of the larger groups of REVIGO linked GO terms of CC and MF processes for the ELT.

CC-Processes	Linked GO terms	MF-Processes	Linked GO terms
Extracellular region part	GO:0044421, GO:0005578, GO:0034361, GO:0005615, GO:0031012, GO:0005576.	**Molecular binding.**	GO:0001948, GO:0030674, GO:0003779, GO:0008092, GO:0008289, GO:0005488, GO:0005509, GO:0005543.
Basolateral plasma membrane.	GO:0016323, GO:0044459, GO:0005581, GO:0043234.	**Structural molecule activity.**	GO:0003777, GO:0008330, GO:0016817, GO:0017111, GO:0008233, GO:0003774, GO:0008237, GO:0004222, GO:0005198, GO:0005200, GO:0005201.
Cytoskeleton.	GO:0005856, GO:0043005, GO:0044463, GO:0005929, GO:0044449, GO:0044441, GO:0043292, GO:0043228, GO:0044430, GO:0043232, GO:0014069, GO:0001533, GO:0042995, GO:0070161, GO:0019861, GO:0030315.	**Kinase, hydrolase and transferase activity.**	GO:0008603, GO:0019207, GO:0047961, GO:0019205, GO:0016787.

Supplemental Table 3. Some of the larger groups of REVIGO linked GO terms of CC and MF processes for the Dams.

CC-Processes	Linked GO terms	MF-Processes	Linked GO terms
Chromosome, centromeric region and protein DNA interaction.	GO:0005694, GO:0000775, GO:0032133, GO:0030141, GO:0019814, GO:0032993.	**Protein activity.**	GO:0004144, GO:0003924, GO:0060089, GO:0004872, GO:0004869, GO:0004950, GO:0004896, GO:0004871, GO:0015144, GO:0005384, GO:0015291, GO:0015294.
Cell surface.	GO:0009897, GO:0005887, GO:0005886, GO:0044459, GO:0009986.	**Cytokine activity.**	GO:0008009, GO:0019956, GO:0005125, GO:0019955, GO:0001664, GO:0043515, GO:0005102, GO:0005070.
Extracellular space.	GO:0044421, GO:0044421, GO:0005615, GO:0042571, GO:0005576.	**Binding.**	GO:0001872, GO:0003823, GO:0005488, GO:0060090, GO:0030246, GO:0005515, GO:0001871.

Supplemental Table 4. Some of the larger groups of REVIGO linked GO terms of CC and MF processes for the Offspring.

CC-Processes	Linked GO terms	MF-Processes	Linked GO terms
Mitochondrion.	GO:0015629, GO:0032432 GO:0042641, GO:0016327, GO:0005938, GO:0044448, GO:0005832, GO:0044427, GO:0005694, GO:0005905, GO:0005581, GO:0043292, GO:0005737, GO:0016023, GO:0000932, GO:0044444, GO:0009898, GO:0005856, GO:0005829, GO:0044445, GO:0022626, GO:0022627, GO:0030286, GO:0005793, GO:0000178, GO:0005794, GO:0000792, GO:0043232, GO:0043229, GO:0044424, GO:0005770, GO:0005811, GO:0030117, GO:0034708, GO:0042613, GO:0042611, GO:0042579, GO:0044455, GO:0044429, GO:0005746, GO:0005739, GO:0032982, GO:0016460, GO:0043228, GO:0005720, GO:0031981, GO:0044428, GO:0005634, GO:0000313, GO:0031090, GO:0044422, GO:0048471, GO:0005777, GO:0046930, GO:0000502, GO:0005839, GO:0043234, GO:0045259, GO:0016469, GO:0030529, GO:0030880, GO:0030017, GO:0000803, GO:0015935, GO:0001725, GO:0042825, GO:0000151, GO:0005774, GO:0005773, GO:0031982.	**Cytoskeletal protein binding**	GO:0003779, GO:0003785, GO:0008013, GO:0005516, GO:0051087, GO:0008092, GO:0019899, GO:0051020, GO:0051427, GO:0042802, GO:0043560, GO:0005178, GO:0035257, GO:0048407, GO:0030674, GO:0032403, GO:0046983, GO:0019904, GO:0046982, GO:0017048, GO:0017124, GO:0046332, GO:0008134, GO:0051082.
		Motor activity	GO:0004559, GO:0042623, GO:0003689, GO:0004521, GO:0016817, GO:0016820, GO:0016788, GO:0015923, GO:0000146, GO:0003774, GO:0016791, GO:0034595, GO:0042578, GO:0033170, GO:0004540, GO:0004298, GO:0070003.
		Ubiquitin-protein transferase activity.	GO:0019200, GO:0003899, GO:0004364, GO:0004402, GO:0042054, GO:0004468, GO:0016278, GO:0004709, GO:0008168, GO:0008080, GO:0019205, GO:0004550, GO:0016773, GO:0016776, GO:0004672, GO:0008276, GO:0004674, GO:0004713, GO:0034062, GO:0016765, GO:0016741, GO:0016772, GO:0004714, GO:0004842.
Cell projection part.	GO:0044463, GO:0005929, GO:0030027, GO:0043005.	**Zinc ion binding.**	GO:0005524, GO:0043169, GO:0019992, GO:0005525, GO:0019001, GO:0043167, GO:0000287, GO:0030145, GO:0046872, GO:0003676, GO:0035091, GO:0017076, GO:0046914, GO:0008270.
Basement membrane.	GO:0005604, GO:0030935.		